PLOS ONE

Mendelian randomization analysis of plasma levels of CD209 and MICB proteins and the risk of varicose veins of lower extremities

Alexandra S. Shadrina[1☯]*, **Elizaveta E. Elgaeva**[1,2☯], **Ian B. Stanaway**[3], **Gail P. Jarvik**[4], **Bahram Namjou**[5], **Wei-Qi Wei**[6], **Joe Glessner**[7], **Hakon Hakonarson**[7], **Pradeep Suri**[8,9,10,11], **Yakov A. Tsepilov**[1,2]*

1 Laboratory of Recombination and Segregation Analysis, Institute of Cytology and Genetics, Novosibirsk, Russia, 2 Department of Natural Sciences, Novosibirsk State University, Novosibirsk, Russia, 3 Division of Nephrology, Department of Medicine, Kidney Research Institute, University of Washington, Seattle, Washington, United States of America, 4 Department of Medicine (Medical Genetics), University of Washington Medical Center, Seattle, Washington, United States of America, 5 Department of Pediatrics, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio, United States of America, 6 Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, Tennessee, United States of America, 7 Department of Pediatrics, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, United States of America, 8 Division of Rehabilitation Care Services, VA Puget Sound Health Care System, Seattle, Washington, United States of America, 9 Seattle Epidemiologic Research and Information Center (ERIC), Department of Veterans Affairs Office of Research and Development, Seattle, Washington, United States of America, 10 Department of Rehabilitation Medicine, University of Washington, Seattle, Washington, United States of America, 11 Clinical Learning, Evidence, and Research (CLEAR) Center, University of Washington, Seattle, Washington, United States of America

☯ These authors contributed equally to this work.
* weiner.alexserg@gmail.com (ASS); tsepilov@bionet.nsc.ru (YAT)

Abstract

Varicose veins of lower extremities (VVs) are a highly prevalent condition, the pathogenesis of which is still not fully elucidated. Mendelian randomization (MR) can provide useful preliminary information on the traits that are potentially causally related to the disease. The aim of the present study is to replicate the effects of the plasma levels of MHC class I polypeptide-related sequence B (MICB) and cluster of differentiation 209 (CD209) proteins reported in a previous hypothesis-free MR study. We conducted MR analysis using a fixed effects inverse-variance weighted meta-analysis of Wald ratios method. For MICB and CD209, we used data from a large-scale genome-wide association study (GWAS) for plasma protein levels (N = 3,301). For VVs, we used GWAS data obtained in the FinnGen project (N = 128,698), the eMERGE network (phase 3, N = 48,429), and the UK Biobank data available in the Gene ATLAS (N = 452,264). The data used in the study were obtained in individuals of European descent. The results for MICB did not pass criteria for statistical significance and replication. The results for CD209 passed all statistical significance thresholds, indicating that the genetically predicted increase in CD209 level is associated with increased risk of VVs (β_{MR} (SE) = 0.07 (0.01), OR (95% CI) = 1.08 (1.05–1.10), P-value = 5.9×10^{-11} in the meta-analysis of three cohorts). Our findings provide further support that CD209 can potentially be involved in VVs. In future studies, independent validation of our results using data from more powerful GWASs for CD209 measured by different methods would be beneficial.

Introduction

Varicose veins of lower extremities (VVs) are a very common vascular disease with a prevalence of over 25–30% in many countries [1–4]. Despite intensive research efforts, the precise mechanisms underlying the development of this condition remain unclear [5–7]. According to current understanding, VVs can be defined as a complex disease with multifactorial pathogenesis which may result from the combined action of a number of factors (genetic, lifestyle, hemodynamic, cellular/extracellular, etc.) [5,8]. Pharmacological treatment of VVs is limited to venoactive medications and several other drugs that can reduce the symptoms of chronic venous disease as well as provide a therapeutic benefit for patients with venous leg ulcers [9,10]. However, developing a drug that can prevent VVs formation, recurrence or progression is still challenging.

Mendelian randomization (MR) is a research method to infer potentially causal relationships between phenotypes. This method uses genetic variants associated with "exposure" phenotypes as naturally occurring "randomizations" [11–13]. When applied to molecular traits (e.g., levels of circulating proteins, lipoproteins, metabolites) or other complex phenotypes (e.g., blood pressure) as "exposure" phenotypes and clinical conditions/diseases as "outcome" phenotypes, MR can be used to propose causative factors, to suggest molecular or physiological pathways contributing to disease, and to discover or prioritize potential targets for pharmacological intervention [14,15].

In our recent study, we applied MR to perform a hypothesis-free search for potentially causal relationships between a broad range of phenotypes and VVs [16]. The "exposure" phenotypes included levels of proteins measured by an aptamer-based affinity proteomics platform (SOMAscan) in blood plasma samples of 1,000 individuals [17]. The "outcome" phenotype was the diagnosis of VVs, and the study sample included 408,455 UK Biobank participants [18]. Our study identified two protein traits that passed all the statistical significance thresholds set in our MR and sensitivity analyses. These traits were plasma levels of MHC class I polypeptide-related sequence B (MICB) protein and cluster of differentiation 209 (CD209) antigen (also known as DC-SIGN–dendritic cell-specific intercellular adhesion molecule 3-grabbing non-integrin). As found by MR, a genetically predicted increase in plasma levels of MICB and CD209 was associated with the presence of VVs, so we concluded that these proteins could potentially be involved in VVs pathogenesis [16]. Both MICB and CD209 are involved in the immune system, and their biological roles are well characterized [19–21]. However, an extensive review of the literature did not allow us to formulate any clear hypothesis on how these proteins may influence the development of VVs. Moreover, MICB and CD209 primarily act as cell surface molecules, while our MR analyzes suggested the effects of circulating (secreted or shed) forms. We raised the question of whether our results could be false-positive findings, or whether we found members of yet undiscovered VVs-related pathways. This question can be answered by conducting experimental research. Before undertaking complex and expensive experimental studies, an optimal strategy would be to perform *in silico* replication using independent datasets. Replication reduces the likelihood that the observed effects are chance findings or analysis artifacts. Conversely, when results are not reproduced in independent *in silico* analyzes, this may indicate that further experimental research will not be fruitful. Thus, the aim of the present study was to replicate the MR results for MICB and CD209 levels as an "exposure" and VVs as an "outcome" phenotype in cohorts different from those used in our previous study [16].

Materials and methods

Datasets

SOMAscan data for MICB and CD209 levels. Genetic association data for plasma levels of MICB and CD209 proteins were obtained from a genetic atlas of the human plasma

proteome (Sun et al. [22]). In that study, relative concentrations of plasma proteins were measured using an expanded version of aptamer-based multiplex protein assay (SOMAscan assay with modified aptamers). The study sample included 3,301 healthy blood donors from the INTERVAL study (recruited in England) genotyped using the Affymetrix Axiom UK Biobank genotyping array [23,24]. Genetic associations were tested by simple linear regression (additive genetic model). Before association testing, relative protein abundances were natural log-transformed and adjusted in a linear regression for age, sex, duration between blood draw and processing and the first three principal components (PCs) of ancestry from multi-dimensional scaling. Then protein residuals were rank-inverse normalized and used as phenotypes in association analysis [22].

FinnGen data for VVs. We downloaded genome-wide association study (GWAS) summary statistics for "Varicose veins (I9_VARICVE)" phenotype from the FinnGen research project website (https://finngen.gitbook.io/documentation/v/r3/data-download; data freeze 3). The FinnGen study combines genotype information of samples collected by a network of Finnish biobanks with digital health record data from Finnish national health registries. For VVs phenotype, the case group included 11,006 subjects with International Classification of Diseases, Tenth Revision (ICD-10) code I83 or ICD-8/9 code 454 ("Varicose veins of lower extremities"; 8,554 women and 2,452 men). The control group included 117,692 individuals without codes related to diseases of veins, lymphatic vessels, and lymph nodes (phlebitis and thrombophlebitis; deep vein thrombosis; portal vein thrombosis; other embolism and thrombosis; oesophageal varices; varicose veins (of lower extremities); varicose veins of other sites; other disorders of veins; nonspesific lymphadenitis; other noninfective disorders of lymphatic vessels and lymph nodes; full lists of codes defining each disease endpoint are available on the website: https://www.finngen.fi/en/researchers/clinical-endpoints). Subjects were genotyped with Illumina (Illumina Inc., USA) and Affymetrix arrays (Thermo Fisher Scientific, USA). Genetic associations were tested using a mixed model logistic regression (SAIGE, Scalable and Accurate Implementation of GEneralized mixed model method [25] which accounts for unbalanced case-control ratios and sample relatedness) with the following covariates included in the model: sex, age, 10 PCs, genotyping batch. Further information on data analysis can be found on the FinnGen study website (https://finngen.gitbook.io/documentation/).

eMERGE data for VVs. We analyzed data from the Electronic Medical Records and Genomics (eMERGE) network, phase 3 [26]. eMERGE is a network of medical centers in the United States (US) with electronic health record (EHR) data linked to biorepository samples and genomic data [26]. The network was supported by funding from the US National Institutes of Health; eMERGE3 involved nine non-pediatric study sites (Columbia University Health Sciences, Geisinger Health, Partners Healthcare/Harvard University, Kaiser Permanente Washington/University of Washington, Mayo Clinic, Marshfield Clinic, Mt. Sinai Health System, Northwestern University, and Vanderbilt University). Further details regarding genotyping and phenotyping in eMERGE3 have been previously reported [26]. Ancestry was determined by the intersection of self-reported race and principal component analysis (PCA)-based ancestry; analyses were restricted to adults of European ancestry who had at least 1 year of EHR data. Cases and controls were defined using phecodes (https://phewascatalog. org/). Specifically, longitudinal EHR data consisting of ICD-9 and ICD-10 codes were used to identify Phecode 454 ("Varicose veins"). Cases were defined as those with 1 or more instances of Phecode 454 (n = 5,800), and controls had no instances of Phecode 454 (n = 42,629). Genotyping was conducted using Illumina and Affymetrix arrays in 83 batches across the participating sites, with imputation of single nucleotide variants (SNVs) performed using guidelines from the Michigan Imputation Server [27,28] and the Haplotype Reference Consortium (HRC) release 1.1 genome build 37 (hg19) reference panel [29]. Third-degree relatives and

closer relatives were excluded from the analysis to account for interrelatedness. Logistic regression of imputed SNVs with an additive genotype model was conducted in R using the glm() function, adjusting for sex, age, site-specific characteristics, and ancestry PCs 1 to 10. Filters were applied for minor allele frequency (MAF) < 0.005, imputation r^2 < 0.3, deviation from Hardy-Weinberg equilibrium (HWE) P-value < 10^{-6}, genotyping call rate < 0.98, and individual call rate < 0.98.

UK Biobank data for VVs. We used genetic association data of UK Biobank study participants available in the Gene ATLAS database (http://geneatlas.roslin.ed.ac.uk/; the second release; data were downloaded in January 2020). Details of the Gene ATLAS study are described in [30]. Details of the UK Biobank study are described in [31–33]. The study cohort was comprised of 452,264 British individuals of European descent. The case group included 12,021 individuals who had ICD-10 code I83 ("Varicose veins of lower extremities") in their medical records, and the control group included 440,243 subjects without this code. Study participants were genotyped with the Affymetrix UK BiLEVE and the Affymetrix UK Biobank Axiom arrays. Associations were tested using a linear mixed model (LMM) method (that allows adjusting for the effect of relatedness), and adjustment was performed for sex, array batch, UK Biobank Assessment Center, age, age^2, and the leading 20 genomic PCs as computed by UK Biobank. The polygenic effect was corrected using a leave-one-chromosome-out (LOCO) approach [30,34]. In our study, we converted the linear regression estimates into the approximate logistic regression estimates. Standard errors (SE) were estimated as $SE = 1/\sqrt{varG \times N \times pr \times (1 - pr)}$, where $varG$ is the variance of genotype, N–sample size, pr–VVs prevalence in the UK Biobank cohort. The variance of genotype was estimated under the assumption of Hardy-Weinberg Equilibrium as $varG = 2 \times p \times (1-p)$, where p is allele frequency. P-values were converted into Z-values, and logistic regression effects (β) were calculated as Z/SE.

Ethics statement

The human plasma proteome study was approved by the National Research Ethics Service (11/EE/0538) [22]. The UK Biobank study was approved by the North West–Haydock Research Ethics Committee (REC reference: 11/NW/0382) [32]. The FinnGen study was approved by the Ethics Committee of the Helsinki and Uusimaa Hospital District (Nr HUS/990/2017). Human subjects approvals were obtained at each participating site as part of eMERGE3. All participants of these studies completed written informed consent.

Two-sample Mendelian randomization

Mendelian randomization analysis of potential causal relationships between CD209 and MICB plasma levels and VVs was performed using a fixed effects inverse-variance weighted meta-analysis of Wald ratios (IVW) approach as previously described by the MR-Base collaboration [35]. MR was conducted using the 'MR-Base' R package ('mr()' and 'mr_report()' functions).

Instrumental variables (single nucleotide polymorphisms, SNPs) were selected from the largest available blood plasma proteome GWAS (N = 3,301 individuals) conducted by Sun et al. [22]. We required the selected SNPs (***i***) to be robustly associated with the exposure trait; (***ii***) to be not in linkage disequilibrium (LD) with each other; (***iii***) to have MAF $\geq$ 0.05. The first (***i***) requirement was fulfilled by selecting only those SNPs that are associated with CD209/MICB levels at a genome-wide significance level of P-value < 5×10^{-8} in two datasets: in Sun et al. dataset [22] and in the blood plasma proteome GWAS conducted by Suhre et al. (N = 1,000 individuals living in southern Germany) [17]. The latter dataset was used as an "exposure" GWAS in our previous MR study [16]. Besides this, we required the selected SNPs

to have the same direction of effect in both these "exposure" GWASs, so that their association with CD209/MICB can be considered as replicated. The second (***ii***) requirement was met by selecting only one representative SNP per LD region by conducting the iterative LD clumping procedure using PLINK 1.9 software [36,37] (https://www.cog-genomics.org/plink2). Thus, our protocol of instrumental variable (IV) selection involved the following steps: first, we obtained overlapping SNPs between Sun et al. [22] and Suhre et al. [17] datasets; second, we excluded SNPs with MAF < 0.05; third, we selected genome-wide significant and replicated (see above) SNPs; fourth, we performed clumping procedure using the PLINK '—clump' function with a 10,000 kb physical distance, *P*-value $< 5 \times 10^{-8}$ significance threshold, and $r^2 >$ 0.001 LD threshold (parameters recommended by the MR-Base, http://app.mrbase.org/). Clumping was performed using Suhre et al. [17] dataset. LD was calculated using 1000 Genomes phase 3 version 5 data for European-ancestry individuals (N = 503). A manual on the clumping procedure can be found at https://zzz.bwh.harvard.edu/plink/clump.shtml. The resulting set of SNPs with association data for CD209/MICB plasma levels is provided in S1 Table.

With these SNPs, we performed three MR analyses: using FinnGen, eMERGE, and Gene ATLAS second release data as an "outcome" GWAS and Sun et al. proteome data [22] as an "exposure" GWAS. Since rs505922 is absent in the FinnGen cohort data, we used a proxy SNP rs576123 for this dataset which is in high LD with rs505922: $r^2 = 0.97$, D' = 0.99 in the FinnGen cohort; $r^2 = 1.00$, D' = 1.00 in Finnish population according to LDlink (https://analysistools.cancer.gov/LDlink/?tab=ldpair; allele rs505922 C is correlated with allele rs576123 C). MR results for FinnGen and eMERGE datasets were meta-analyzed. Finally, we conducted a meta-analysis summarizing all three MR tests. Meta-analysis was performed using a fixed effects IVW approach, and heterogeneity was assessed using the Cochran's Q test. In addition to heterogeneity assessment, we applied a two-sample t-test to compare the MR beta (β_{MR}) values and their standard errors (SEs) obtained using each pair of datasets. The statistical significance threshold for the Q test and t-tests was set at *P*-value < 0.05.

The criteria for statistically significant and replicated results in the present study were as follows: (1a) β_{MR} sign obtained using Gene ATLAS second release data is the same as β_{MR} sign obtained in our earlier study [16] (S2 Table); (1b) the *P*-value in this MR analysis is less than 1.1×10^{-5} (the threshold used in our previous work [16]); (2a) β_{MR} sign in the meta-analysis of MR results obtained using FinnGen and eMERGE data is the same as β_{MR} sign obtained in our earlier study [16] (S2 Table); (2b) the *P*-value in this meta-analysis is less than 0.025 (0.05/2) (Fig 1).

Sensitivity analyses

For CD209, we performed sensitivity tests. First, for each separate IV, we conducted MR analyses for all datasets, meta-analyzed MR results with heterogeneity assessment and then compared β_{MR} values and their SEs between each pair of datasets and between a pair of IVs using a two-sample t-test.

Second, we performed additional MR analyses considering IVs suggestively associated (*P*-value $< 5 \times 10^{-7}$) with CD209 level. This was done to increase the number of IVs and perform the tests that could not be performed with a limited number of IVs in the main analysis. For these sensitivity analyses, we used genetic association data for CD209 plasma level from Sun et al. [22] dataset as an "exposure" GWAS and Gene ATLAS second release data for VVs (as the largest dataset for VVs) as an "outcome" GWAS.

IVs were selected using the clumping procedure implemented in PLINK 1.9 software [36,37] with the same settings as described above for the main analysis, except for the statistical
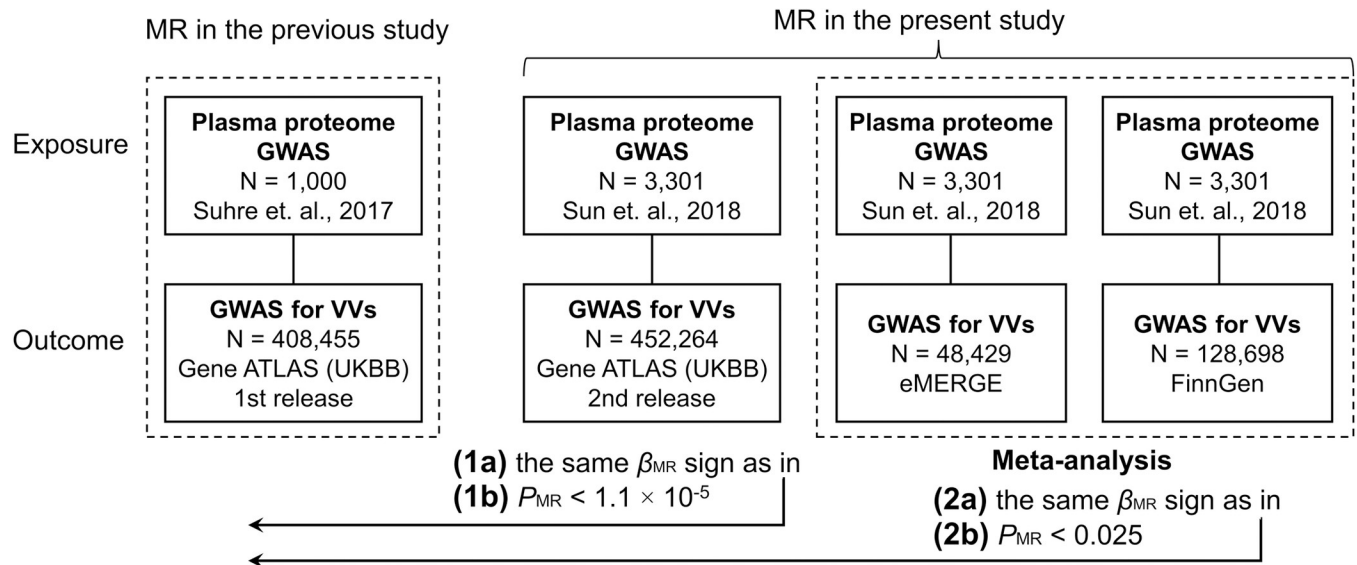
Fig 1. Criteria for statistically significant and replicated results used in the present study. GWAS, genome-wide association study; MR, Mendelian randomization; *P*, *P*-value; UKBB, UK Biobank; VVs, varicose veins of lower extremities.

significance threshold of *P*-value = 5×10^{-7}. Since there were no additional IVs under the relaxed threshold in Suhre at al. [17] GWAS for CD209, we performed clumping using Sun et al. [22] dataset. The list of five selected IVs suggestively associated with CD209 is provided in S3 Table. For rs151212242 variant that is absent in Gene ATLAS data, we used a proxy SNP rs4804224 (r^2 = 1.0, D' = 1.0 in European population; allele rs151212242 C is correlated with allele rs4804224 C).

MR analyses were conducted using five MR methods (MR-Egger, Weighted median, IVW, Simple mode, and Weighted mode [35]) integrated into the 'TwoSampleMR' version 0.5.5 R package. Summary statistics for IVs obtained from the exposure and outcome datasets was harmonized using the 'harmonise_data()' function, and MR tests were performed with the 'mr ()' and 'mr_report()' functions. Besides the MR tests, the tests for heterogeneity, the test for directional horizontal pleiotropy, and the test identifying the correct direction of effect (Steiger test) embedded into 'TwoSampleMR' package were carried out. The presence of horizontal pleiotropy was additionally assessed using the 'mr_presso()' function of the 'MR-PRESSO' version 1.0 R package [38].

Finally, we used the MR–Robust Adjusted Profile Score (MR-RAPS) approach implemented into the 'mr.raps' version 0.4 R package ('mr.raps.mle.all()' function) to account for the potential bias introduced by selecting weaker IVs [39].

Results

The results of the IVW MR analysis for MICB and CD209 plasma levels performed using different datasets for the outcome VVs trait are presented in Tables 1 and S4.

MR analysis for CD209

CD209 passed all the statistical significance thresholds set in our study with the resulting β_{MR} (SE) of 0.07 (0.01), odds ratio (OR) of 1.08 with 95% confidence interval (95% CI) of 1.05–1.10, and *P*-value of 5.9×10^{-11} in the meta-analysis of MR results for FinnGen, eMERGE, and Gene ATLAS (second release) cohorts. Thus, our results confirmed our previous observations

that the genetically predicted increase in CD209 plasma level is associated with VVs suggesting a causal effect of CD209 on the development of this pathology.

MR analysis for CD209 was conducted using two IVs that met our strict selection criteria (S1 Table; both IVs were the same as those used in our previous study [16]). One of them, rs505922, is located in the gene encoding ABO blood group glycosyltransferases. The variant rs505922 is in LD with rs8176719, a key SNP responsible for blood group O status [40] ($r^2 = 0.87$, D' = 0.99 in European-ancestry populations according to LDlink). Besides this, rs505922 is in LD with rs507666 ($r^2 = 0.39$, D' = 1.00), which is one of the top SNPs associated with VVs in the "23andMe" GWAS [41] and subsequently replicated in our previous work [42]. According to our estimation, rs505922 explains as much as 40% of the variability in CD209 plasma level [16]. The second IV used in MR was rs8106657. This SNP is located nearly 17 kb from the CD209 gene representing a *cis*-protein quantitative trait locus (pQTL). To assess the impact of each IV on MR results, we repeated a full set of tests with each IV separately (Tables 2 and S5). The results of rs505922-based MR were statistically significant and consistent across all datasets. However, MR with rs8106657 alone did not produce any statistically significant findings, except for a nominally significant result for the Gene ATLAS dataset ($P_{MR} = 0.02$), and β_{MR} values were generally less than those obtained in MR with rs505922 alone. For the eMERGE dataset, β_{MR} in rs8106657-based MR had a negative sign as opposed to the remaining MR analyses, and for the Gene ATLAS dataset, β_{MR} in rs8106657-based MR was 0.10 while for the eMERGE and FinnGen datasets it was close to zero. This prompted us to speculate that the results of the primary MR performed with both IVs could be driven by the effect of rs505922. To check this, we compared β_{MR} values (and their SEs) obtained in rs8106657- and rs505922-based MR analyzes assuming that statistically significant differences between these values would indicate a difference in effects. Nevertheless, the two-sample t-tests did not reveal any statistically significant differences between the pairs of β_{MR} (and their SEs) for rs8106657 and rs505922 (Tables 2 and S5), neither revealed any differences in β_{MR} (and their SEs) between pairs of datasets in the single-IV MR analyses (Tables 2 and S5) and in the primary MR based on both IVs (Tables 1 and S4). Thus, our tests did not provide evidence that rs505922 is fully responsible for the overall effect revealed in our MR analysis.

Further we performed a sensitivity analysis using an extended set of IVs suggestively associated with CD209 plasma level (S3 Table) and five MR methods: MR-Egger, Weighted median, IVW, Simple mode, and Weighted mode. The results were concordant with those obtained in the main IVW MR analysis (S6 Table). Heterogeneity tests provided no evidence for statistically significant heterogeneity in causal effects amongst instruments; horizontal pleiotropy tests (the test based on the MR-Egger regression intercept and the MR-PRESSO global test) showed no evidence for directional horizontal pleiotropy; and the Steiger test did not identify the wrong direction of causality (S6 Table). Finally, we used the MR-RAPS approach to account for the potential bias introduced by weaker IVs. The results of the MR-RAPS analysis were consistent with those of the main analysis (S7 Table) indicating that the detected effect is robust regardless of the number and strength of the selected IVs.

MR analysis for MICB

For MICB, MR was performed using a single IV rs3094005, which is a *cis*- pQTL located in the *MICB* gene. The results of MR for this protein did not meet the criteria for statistically significant and replicated results set in our study. Firstly, the *P*-value in the MR analysis performed using Gene ATLAS second release data was higher than the threshold used in our previous study based on the first release of the Gene ATLAS database [16] (3.3×10^{-5} vs. 1.1×10^{-5}). Secondly, the *P*-value in the meta-analysis of MR results for FinnGen and eMERGE cohorts

Table 1. Results of the IVW Mendelian randomization analysis of the effects of MICB and CD209 plasma levels on VVs.

	Gene ATLAS (second release)[a] (N = 452,264)		eMERGE (N = 48,429)		FinnGen (N = 128,698)		Meta-analysis eMERGE + FinnGen (N = 177,127)		Meta-analysis Gene ATLAS[a] + eMERGE + FinnGen (N = 629,391)				Two-sample t-test (comparison of β_{MR} and their SEs)		
N_{IV}[b]	β_{MR} (SE)	P_{MR}	β_{MR} (SE)	P_{MR}	β_{MR} (SE)	P_{MR}	β_{MR} (SE)	P_{MR}	β_{MR} (SE)	OR (95% CI)	P_{MR}	P_{het}[c]	P_{GE}[d]	P_{FE}[e]	P_{GF}[f]
CD209 vs. VVs															
2	0.08 (0.02)	2.0e-7	0.06 (0.03)	0.04	0.07 (0.02)	5.7e-4	0.07 (0.02)	6.1e-5	0.07 (0.01)	1.08 (1.05–1.10)	5.9e-11	0.80	0.50	0.68	0.80
MICB vs. VVs															
1	0.14 (0.03)	3.3e-5	-0.09 (0.10)	0.38	0.12 (0.04)	5.4e-3	0.09 (0.04)	0.03	0.12 (0.03)	1.13 (1.07–1.18)	4.7e-6	0.09	0.03	0.05	0.75

Sources of data for the outcome trait (VVs) are indicated in the first raw.

More precise numerical data (with larger number of decimal places) and all OR and 95% CI values are provided in S4 Table.

CI, confidence interval; IV, instrumental variable; MR, Mendelian randomization; OR, odds ratio; P, P-value; SE, standard error.

[a] UK Biobank data available in the Gene ATLAS database (second release) [30].

[b] Number of instrumental variables used in the analysis (listed in S1 Table).

[c] P-value for heterogeneity assessed using Cochran's Q test.

[d] P-value for comparison of β_{MR} values and their SEs between the Gene ATLAS and eMERGE datasets using a two-sample t-test.

[e] P-value for comparison of β_{MR} values and their SEs between FinnGen and eMERGE datasets using a two-sample t-test.

[f] P-value for comparison of β_{MR} values and their SEs between the Gene ATLAS and FinnGen datasets using a two-sample t-test.

https://doi.org/10.1371/journal.pone.0268725.t001

was 0.03, which is higher that the Bonferroni-corrected threshold of 0.025 (Tables 1 and S4). Besides this, β_{MR} obtained using the eMERGE dataset had the opposite sign to β_{MR} obtained in the Gene ATLAS- and FinnGen-based analyses (-0.09 vs. 0.14 and 0.12, respectively). Hence, although the P-value in the meta-analysis of the data for the three cohorts was 4.7×10^{-6}, we do not consider the association between MICB plasma level and VVs as replicated in our study.

Table 2. Results of IVW Mendelian randomization analysis of the effect of CD209 plasma levels on VVs considering each selected instrumental variable separately.

	Gene ATLAS (second release)[a] (N = 452,264)		eMERGE (N = 48,429)		FinnGen (N = 128,698)		Meta-analysis eMERGE + FinnGen (N = 177,127)		Meta-analysis Gene ATLAS[a] + eMERGE + FinnGen (N = 629,391)				Two-sample t-test (comparison of β_{MR} and their SEs)		
IV[b]	β_{MR} (SE)	P_{MR}	β_{MR} (SE)	P_{MR}	β_{MR} (SE)	P_{MR}	β_{MR} (SE)	P_{MR}	β_{MR} (SE)	OR (95% CI)	P_{MR}	P_{het}[c]	P_{GE}[d]	P_{FE}[e]	P_{GF}[f]
rs505922[g]	0.08 (0.02)	2.6e-6	0.07 (0.03)	0.01	0.08 (0.02)	4.2e-5	0.08 (0.02)	1.5e-6	0.08 (0.01)	1.08 (1.06–1.10)	1.7e-11	0.94	0.80	0.72	0.88
rs8106657	0.10 (0.04)	0.02	-0.01 (0.07)	0.85	0.02 (0.05)	0.76	0.01 (0.04)	0.89	0.05 (0.03)	1.05 (0.99–1.11)	0.09	0.30	0.18	0.73	0.24
P_{iv}[h] t-test	0.68		0.27		0.25		0.11		0.40						

Sources of data for the outcome trait (VVs) are indicated in the first raw.

Extended data including all OR and 95% CI values are provided in S5 Table.

CI, confidence interval; IV, instrumental variable; MR, Mendelian randomization; OR, odds ratio; P, P-value; SE, standard error.

[a] UK Biobank data available in the Gene ATLAS database (second release) [30].

[b] Instrumental variables used in the analysis (more details are provided in S1 Table).

[c] P-value for heterogeneity assessed using Cochran's Q test.

[d] P-value for comparison of β_{MR} values and their SEs between the Gene ATLAS and eMERGE datasets using a two-sample t-test.

[e] P-value for comparison of β_{MR} values and their SEs between FinnGen and eMERGE datasets using a two-sample t-test.

[f] P-value for comparison of β_{MR} values and their SEs between the Gene ATLAS and FinnGen datasets using a two-sample t-test.

[g] rs576123 was used as a proxy for rs505922 in the MR analysis for the FinnGen cohort ($r^2 = 0.97$, D' = 0.99).

[h] P-value for comparison of β_{MR} values and their SEs between IVs using a two-sample t-test.

https://doi.org/10.1371/journal.pone.0268725.t002

Discussion

In the present study, we used data from three large-scale GWAS for VVs and the largest available GWAS for human plasma proteome to investigate the relationship between plasma levels of MICB and CD209 proteins and VVs via Mendelian randomization. Our results confirmed the association of the genetically predicted increase in CD209 level with VVs revealed in our previous study [16]. The results for MICB did not pass the pre-defined statistical significance criteria (Fig 1), indicating that our previous observation was a false positive.

CD209 (DC-SIGN) is a C-type lectin transmembrane receptor protein primarily expressed by dendritic cells (DC). CD209 acts as a cell adhesion molecule and plays an important role in DC functioning, including interaction with endothelial cells, T-cells, and neutrophils. Besides this, CD209 mediates recognition of a wide variety of pathogens (viruses, bacteria, fungi, parasites) and is involved in their capture and internalization [20,21]. Plazolles et al. [43] demonstrated the presence of a full-length soluble secreted form of CD209 (sDC-SIGN) in several human body fluids such as serum, joint fluids, and bronchoalveolar lavages and showed that its expression appears to be up-regulated upon inflammation. The functional role of the soluble CD209 form and the mechanism of its secretion remain largely unknown, although recent studies have linked changes in its serum level to non-Hodgkin lymphoma [44] and colon cancer [45] and proposed that sDC-SIGN could enhance cytomegalovirus infection [43]. In our MR analyses (both in this and in the previous study [16]), we used data for CD209 determined in plasma using the SOMAscan assay which measures both extracellular and intracellular proteins (including soluble domains of membrane-associated proteins) with a bias towards secreted proteins [22]. Thus, the association with VVs is likely to be shown for the soluble rather than the transmembrane form of CD209. Given the lack of knowledge about the sDC-SIGN function, it is currently challenging to propose a mechanism for its involvement in VVs. However, since the inflammatory response is activated in VVs and is considered part of their pathogenesis [5–8,46,47], we can speculate that this potential link is related to inflammation.

Albeit the association between CD209 and VVs was found in our two studies using independent datasets, the putative causative effect of plasma CD209 level on VVs development has yet to be confirmed in future research. First of all, positive MR results by themselves are insufficient for making a causal claim [48], so *in vitro* and *in vivo* studies are necessary to draw a final conclusion. Besides this, our study has several limitations that must be acknowledged.

The first limitation is a small number of instrumental variables used in the analysis. For CD209, available plasma proteome GWASs [17,22] provide only two genome-wide significant SNPs not in LD with each other (one *cis* and one *trans* pQTL) that can be used as strong IVs in MR. One of them, rs505922, is located in the *ABO* gene in LD with blood group O- and A_1-tagging SNPs. These SNPs exert pleiotropic effects on different human traits [49], including the levels of soluble leukocyte adhesion molecules ICAM-1, E-selectin, and P-selectin [50–52]. Since the presence of CD209-independent effect of rs505922 on the risk of VVs can be hypothesized [42], violation of the 'no horizontal pleiotropy' assumption in MR cannot be ruled out. The inclusion of a larger number of IVs in the MR analysis, on the one hand, would enable performing sensitivity tests, and on the other hand, in theory, can lead to a "dilution" of independent pleiotropic effects relative to associations with the trait of interest [53]. Thus, if more powerful plasma proteome GWASs reveal more CD209-associated SNPs in future, replication of our results would be highly beneficial. To test the general reliability of our findings, we compared β_{MR} values (and their SEs) obtained in single-IV MR analyzes for both SNPs to check whether MR results could be fully driven by rs505922. However, our tests did not confirm this assumption. Next, we performed a sensitivity analysis using five IVs including weaker ones

associated with CD209 at a suggestive level of statistical significance. This allowed us to use more MR methods and perform additional tests. The results of all MR methods and approaches used were concordant with the results of the main analysis, and no evidence was observed for directional horizontal pleiotropy, heterogeneity between IVs, and the wrong direction of causality.

The second limitation is related to the same method used to measure CD209 in both Suhre et al. [17] and Sun et al. [22] plasma proteome GWASs, the first of which was used in our previous hypothesis-free study [16], and the second of which was used in the present study. If we assume that the aptamer-based method provides biased estimates of CD209 plasma levels, our replication study will suffer from the same problems as the primary study.

The third limitation is that our study only included data obtained from European-ancestry individuals. Thus, the results of our study could not be generalizable to other populations.

Fourth, a phenotyping approach based on the extraction of ICD codes from medical records used in the "outcome" GWASs cannot guarantee that all VVs cases have a confirmed diagnosis and all controls do not have VVs. Of note, the prevalence of VVs was nearly 12% in the eMERGE cohort, 8.6% in the FinnGen cohort, and only 2.7% in the UK Biobank cohort, that is much lower than VVs prevalence estimates obtained in many countries of the Western world (generally over 20%) [1–3]. Nevertheless, this limitation may be compensated for by the large size of the analyzed datasets.

The fifth limitation is that for sensitivity analyses considering IVs suggestively associated with CD209 plasma levels, the same dataset was used for both IV selection and MR analyses, which could lead to so-called "selection bias" [39].

Finally, a confounding effect of inflammation can be proposed if the same inflammatory pathways promote the release of soluble CD209 (sDC-SIGN) [43] and affect the pathogenesis of VVs. Association of IVs with such inflammatory factors, if any, can lead to bias in MR analyses [35].

Conclusions

Our study provided further evidence that a genetically predicted increase in plasma CD209 level is associated with the risk of VVs, supporting CD209 as a candidate for future studies of the molecular mechanisms of VVs pathogenesis. An independent *in silico* validation of our MR results using an expanded set of instrumental variables from the GWASs with different methods of CD209 measurement (as they become available) would be beneficial.

Supporting information

S1 Table. Summary statistics for instrumental variables genome-wide significantly associated with CD209/MICB levels in human blood plasma proteome GWASs.
(XLSX)

S2 Table. Results of hypothesis-free 2SMR analysis for CD209 and MICB vs. VVs obtained in our previous study.
(XLSX)

S3 Table. Summary statistics for instrumental variables suggestively associated with CD209 level.
(XLSX)

S4 Table. Results of the IVW Mendelian randomization analysis of the effects of MICB and CD209 plasma levels on VVs.
(XLSX)

S5 Table. Results of the IVW Mendelian randomization analysis of the effect of CD209 plasma levels on VVs considering each selected instrumental variable separately.
(XLSX)

S6 Table. Results of the Mendelian randomization analysis of the effect of CD209 plasma level on VVs considering instrumental variables suggestively associated with CD209 plasma level.
(XLSX)

S7 Table. Results of MR-RAPS analysis of the effect of CD209 plasma levels on VVs considering instrumental variables suggestively associated with CD209 plasma level.
(XLSX)

Acknowledgments

Author Contributions

Conceptualization: Alexandra S. Shadrina, Yakov A. Tsepilov.

Data curation: Elizaveta E. Elgaeva, Yakov A. Tsepilov.

Formal analysis: Elizaveta E. Elgaeva, Ian B. Stanaway, Yakov A. Tsepilov.

Funding acquisition: Yakov A. Tsepilov.

Investigation: Alexandra S. Shadrina, Elizaveta E. Elgaeva, Yakov A. Tsepilov.

Methodology: Alexandra S. Shadrina, Elizaveta E. Elgaeva, Yakov A. Tsepilov.

Resources: Gail P. Jarvik, Bahram Namjou, Wei-Qi Wei, Joe Glessner, Hakon Hakonarson, Pradeep Suri.

Supervision: Yakov A. Tsepilov.

Writing – original draft: Alexandra S. Shadrina.

Writing – review & editing: Elizaveta E. Elgaeva, Ian B. Stanaway, Gail P. Jarvik, Bahram Namjou, Wei-Qi Wei, Joe Glessner, Hakon Hakonarson, Pradeep Suri, Yakov A. Tsepilov.

References

1. Beebe-Dimmer JL, Pfeifer JR, Engle JS, Schottenfeld D. The epidemiology of chronic venous insufficiency and varicose veins. Ann Epidemiol. 2005; 15: 175–184. https://doi.org/10.1016/j.annepidem.2004.05.015 PMID: 15723761

2. Evans CJ, Fowkes FGR, Ruckley C V., Lee AJ. Prevalence of varicose veins and chronic venous insufficiency in men and women in the general population: Edinburgh Vein Study. J Epidemiol Community Health. 1999; 53: 149–153. https://doi.org/10.1136/jech.53.3.149 PMID: 10396491

3. Zolotukhin IA, Seliverstov EI, Shevtsov YN, Avakiants IP, Nikishkov AS, Tatarintsev AM, et al. Prevalence and risk factors for chronic venous disease in the general Russian population. Eur J Vasc Endovasc Surg. 2017; 54: 752–758. https://doi.org/10.1016/j.ejvs.2017.08.033 PMID: 29031868

4. Agarwal V, Agarwal S, Singh A, Nathwani P, Goyal P, Goel S. Prevalence and risk factors of varicose veins, skin trophic changes, and venous symptoms among northern Indian population. Int J Res Med Sci. 2016; 4: 1678–1682.

5. Oklu R, Habito R, Mayr M, Deipolyi AR, Albadawi H, Hesketh R, et al. Pathogenesis of varicose veins. J Vasc Interv Radiol; 2012; 23: 33–39. https://doi.org/10.1016/j.jvir.2011.09.010 PMID: 22030459

6. Jacobs BN, Andraska EA, Obi AT, Wakefield TW. Pathophysiology of varicose veins. J Vasc Surg Venous Lymphat Disord. 2017; 5: 460–467. https://doi.org/10.1016/j.jvsv.2016.12.014 PMID: 28411716

7. Lim CS, Davies AH. Pathogenesis of primary varicose veins. Br J Surg. 2009; 96: 1231–1242. https://doi.org/10.1002/bjs.6798 PMID: 19847861

8. Segiet OA, Brzozowa-Zasada M, Piecuch A, Dudek D, Reichman-Warmusz E, Wojnicz R. Biomolecular mechanisms in varicose veins development. Ann Vasc Surg. 2015; 29: 377–384. https://doi.org/10.1016/j.avsg.2014.10.009 PMID: 25449990

9. Nicolaides AN. The benefits of micronized purified flavonoid fraction (MPFF) throughout the progression of chronic venous disease. Adv Ther. 2020; 37: 1–5. https://doi.org/10.1007/s12325-019-01218-8 PMID: 31970659

10. Carroll BJ, Piazza G, Goldhaber SZ. Sulodexide in venous disease. J Thromb Haemost. 2019; 17: 31–38. https://doi.org/10.1111/jth.14324 PMID: 30394690

11. Davies NM, Holmes M V., Davey Smith G. Reading Mendelian randomisation studies: A guide, glossary, and checklist for clinicians. BMJ. 2018; 362: k601. https://doi.org/10.1136/bmj.k601 PMID: 30002074

12. Pingault JB, O'Reilly PF, Schoeler T, Ploubidis GB, Rijsdijk F, Dudbridge F. Using genetic data to strengthen causal inference in observational research. Nat Rev Genet. 2018; 19: 566–580. https://doi.org/10.1038/s41576-018-0020-3 PMID: 29872216

13. Morrison J, Knoblauch N, Marcus JH, Stephens M, He X. Mendelian randomization accounting for correlated and uncorrelated pleiotropic effects using genome-wide summary statistics. Nat Genet. 2020; 52: 740–747. https://doi.org/10.1038/s41588-020-0631-4 PMID: 32451458

14. Zheng J, Haberland V, Baird D, Walker V, Haycock PC, Hurle MR, et al. Phenome-wide Mendelian randomization mapping the influence of the plasma proteome on complex diseases. Nat Genet. 2020; 52: 1122–1131. https://doi.org/10.1038/s41588-020-0682-6 PMID: 32895551

15. Schmidt AF, Finan C, Gordillo-Marañón M, Asselbergs FW, Freitag DF, Patel RS, et al. Genetic drug target validation using Mendelian randomisation. Nat Commun. 2020; 11: 1–12.

16. Shadrina AS, Sharapov SZ, Shashkova TI, Tsepilov YA. Varicose veins of lower extremities: Insights from the first large-scale genetic study. PLOS Genet. 2019; 15: e1008110. https://doi.org/10.1371/journal.pgen.1008110 PMID: 30998689

17. Suhre K, Arnold M, Bhagwat AM, Cotton RJ, Engelke R, Raffler J, et al. Connecting genetic risk to disease end points through the human blood plasma proteome. Nat Commun. 2017; 8: 14357. https://doi.org/10.1038/ncomms14357 PMID: 28240269

18. Canela-Xandri O, Rawlik K, Tenesa A. An atlas of genetic associations in UK Biobank. bioRxiv: 176834v2 [Preprint]. 2017 [cited 2022 Mar 19]. Available from: https://doi.org/10.1101/176834.

19. Collins RWM. Human MHC class I chain related (MIC) genes: Their biological function and relevance to disease and transplantation. Eur J Immunogenet. 2004; 31: 105–114. https://doi.org/10.1111/j.1365-2370.2004.00457.x PMID: 15182323

20. Geurtsen J, Driessen NN, Appelmelk BJ. Mannose-fucose recognition by DC-SIGN. In: Microbial Glycobiology. Elsevier Inc.; 2010. p. 673–695.

21. García-Vallejo JJ, van Liempt E, da Costa Martins P, Beckers C, van het Hof B, Gringhuis SI, et al. DC-SIGN mediates adhesion and rolling of dendritic cells on primary human umbilical vein endothelial cells through LewisY antigen expressed on ICAM-2. Mol Immunol. 2008; 45: 2359–2369. https://doi.org/10.1016/j.molimm.2007.11.001 PMID: 18155766

22. Sun BB, Maranville JC, Peters JE, Stacey D, Staley JR, Blackshaw J, et al. Genomic atlas of the human plasma proteome. Nature. 2018; 558: 73–79. https://doi.org/10.1038/s41586-018-0175-2 PMID: 29875488

23. Moore C, Sambrook J, Walker M, Tolkien Z, Kaptoge S, Allen D, et al. The INTERVAL trial to determine whether intervals between blood donations can be safely and acceptably decreased to optimise blood supply: Study protocol for a randomised controlled trial. Trials. 2014; 15: 363. https://doi.org/10.1186/1745-6215-15-363 PMID: 25230735

24. Astle WJ, Elding H, Jiang T, Allen D, Ruklisa D, Mann AL, et al. The allelic landscape of human blood cell trait variation and links to common complex disease. Cell. 2016; 167: 1415–1429.e19. https://doi.org/10.1016/j.cell.2016.10.042 PMID: 27863252

25. Zhou W, Nielsen JB, Fritsche LG, Dey R, Gabrielsen ME, Wolford BN, et al. Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. Nat Genet. 2018; 50: 1335–1341. https://doi.org/10.1038/s41588-018-0184-y PMID: 30104761

26. Stanaway IB, Hall TO, Rosenthal EA, Palmer M, Naranbhai V, Knevel R, et al. The eMERGE genotype set of 83,717 subjects imputed to ~40 million variants genome wide and association with the herpes zoster medical record phenotype. Genet Epidemiol. 2019; 43: 63–81. https://doi.org/10.1002/gepi.22167 PMID: 30298529

27. Das S, Forer L, Schönherr S, Sidore C, Locke AE, Kwong A, et al. Next-generation genotype imputation service and methods. Nat Genet. 2016; 48: 1284–1287. https://doi.org/10.1038/ng.3656 PMID: 27571263

28. Loh PR, Danecek P, Palamara PF, Fuchsberger C, Reshef YA, Finucane HK, et al. Reference-based phasing using the Haplotype Reference Consortium panel. Nat Genet. 2016; 48: 1443–1448. https://doi.org/10.1038/ng.3679 PMID: 27694958

29. The Haplotype Reference Consortium. A reference panel of 64,976 haplotypes for genotype imputation. Nat Genet. 2016; 48: 1279–1283. https://doi.org/10.1038/ng.3643 PMID: 27548312

30. Canela-Xandri O, Rawlik K, Tenesa A. An atlas of genetic associations in UK Biobank. Nat Genet. 2018; 50: 1593–1599. https://doi.org/10.1038/s41588-018-0248-z PMID: 30349118

31. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK Biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLOS Med. 2015; 12: e1001779. https://doi.org/10.1371/journal.pmed.1001779 PMID: 25826379

32. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK Biobank resource with deep phenotyping and genomic data. Nature. 2018; 562: 203–209. https://doi.org/10.1038/s41586-018-0579-z PMID: 30305743

33. Fry A, Littlejohns TJ, Sudlow C, Doherty N, Adamska L, Sprosen T, et al. Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. Am J Epidemiol. 2017; 186: 1026–1034. https://doi.org/10.1093/aje/kwx246 PMID: 28641372

34. Yang J, Zaitlen NA, Goddard ME, Visscher PM, Price AL. Advantages and pitfalls in the application of mixed-model association methods. Nat Genet. 2014; 46: 100–106. https://doi.org/10.1038/ng.2876 PMID: 24473328

35. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, et al. The MR-Base platform supports systematic causal inference across the human phenome. Elife. 2018; 7: e34408. https://doi.org/10.7554/eLife.34408 PMID: 29846171

36. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007; 8: 559–575. https://doi.org/10.1086/519795 PMID: 17701901

37. Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: Rising to the challenge of larger and richer datasets. Gigascience. 2015; 4: 7.

38. Verbanck M, Chen CY, Neale B, Do R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. Nat Genet. 2018; 50: 693–698. https://doi.org/10.1038/s41588-018-0099-7 PMID: 29686387

39. Zhao Q, Wang J, Hemani G, Bowden J, Small DS. Statistical inference in two-sample summary-data Mendelian randomization using robust adjusted profile score. Ann Statist. 2020; 48: 1742–1769.

40. Yamamoto FI, Clausen H, White T, Marken J, Hakomori SI. Molecular genetic basis of the histo-blood group ABO system. Nature. 1990; 345: 229–233. https://doi.org/10.1038/345229a0 PMID: 2333095

41. Bell RK, Durand EY, McLean CY, Eriksson N, Tung JY, Hinds D. A large scale genome wide association study of varicose veins in the 23andMe cohort. In: The 64th Annual Meeting of The American Society of Human Genetics, San Diego, California, USA, 18–22 October 2014, paper no. 2082M, p.487. San Diego: ASHG. [cited 2022 Mar 19]. Available from: https://blog.23andme.com/wp-content/uploads/2014/10/Bell_ASHG2014_varicose.pdf.

42. Shadrina A, Tsepilov Y, Smetanina M, Voronina E, Seliverstov E, Ilyukhin E, et al. Polymorphisms of genes involved in inflammation and blood vessel development influence the risk of varicose veins. Clin Genet. 2018; 94: 191–199. https://doi.org/10.1111/cge.13362 PMID: 29660117

43. Plazolles N, Humbert J, Vachot L, Verrier B, Hocke C, Halary F. Pivotal advance: The promotion of soluble DC-SIGN release by inflammatory signals and its enhancement of cytomegalovirus-mediated cis-infection of myeloid dendritic cells. J Leukoc Biol. 2011; 89: 329–342. https://doi.org/10.1189/jlb.0710386 PMID: 20940323

44. Ding D, Chen W, Zhang C, Chen Z, Jiang Y, Yang Z, et al. Low expression of dendritic cell-specific intercellular adhesion molecule-3-grabbing nonintegrin in non-Hodgkin lymphoma and a significant

correlation with β2-microglobulin. Med Oncol. 2014; 31: 1–10. https://doi.org/10.1007/s12032-014-0202-6 PMID: 25182705

45. Jiang Y, Zhang C, Chen K, Chen Z, Sun Z, Zhang Z, et al. The cinical significance of DC-SIGN and DC-SIGNR, which are novel markers expressed in human colon cancer. PLoS One. 2014; 9: e114748. https://doi.org/10.1371/journal.pone.0114748 PMID: 25504222

46. Zolotukhin IA, Porembskaya OY, Smetanina MA, Sazhin AV, Kirienko AI. Varicose veins: On the verge of discovering the cause? Ann Russ Acad Med Sci. 2020; 75: 36–45.

47. Ghaderian SMH, Lindsey NJ, Graham AM, Homer-Vanniasinkam S, Najar RA. Pathogenic mechanisms in varicose vein disease: the role of hypoxia and inflammation. Pathology. 2010; 42: 446–453. https://doi.org/10.3109/00313025.2010.493865 PMID: 20632821

48. Burgess S, O'Donnell CJ, Gill D. Expressing results from a Mendelian randomization analysis: Separating results from inferences. JAMA Cardiol. 2021; 6: 7–8. https://doi.org/10.1001/jamacardio.2020.4317 PMID: 32965465

49. Li S, Schooling CM. A phenome-wide association study of ABO blood groups. BMC Med 2020; 18: 1–11.

50. Paré G, Chasman DI, Kellogg M, Zee RYL, Rifai N, Badola S, et al. Novel association of ABO histo-blood group antigen with soluble ICAM-1: Results of a genome-wide association study of 6,578 women. PLoS Genet. 2008; 4: e1000118. https://doi.org/10.1371/journal.pgen.1000118 PMID: 18604267

51. Paterson AD, Lopes-Virella MF, Waggott D, Boright AP, Hosseini SM, Carter RE, et al. Genome-wide association identifies the ABO blood group as a major locus associated with serum levels of soluble E-selectin. Arterioscler Thromb Vasc Biol. 2009; 29: 1958–1967. https://doi.org/10.1161/ATVBAHA.109.192971 PMID: 19729612

52. Barbalic M, Dupuis J, Dehghan A, Bis JC, Hoogeveen RC, Schnabel RB, et al. Large-scale genomic studies reveal central role of ABO in sP-selectin and sICAM-1 levels. Hum Mol Genet. 2010; 19: 1863–1872. https://doi.org/10.1093/hmg/ddq061 PMID: 20167578

53. Swerdlow DI, Kuchenbaecker KB, Shah S, Sofat R, Holmes MV, White J, et al. Selecting instruments for Mendelian randomization in the wake of genome-wide association studies. Int J Epidemiol. 2016; 45: 1600–1616. https://doi.org/10.1093/ije/dyw088 PMID: 27342221